

הכלי ClustalW

ClustalW – تُستخدم هذه الأداة لمُقارنة تسلسلين أو أكثر، بواسطة طريقة تُسمى التراصف التسلسلي (העלמדת רצפים - Sequence Alignment).



ClustalW – تُستخدم هذه الأداة لمُقارنة تسلسلات من الأحماض الأمينية أو تسلسلات من النوكليوتيدات بواسطة طريقة تُسمى التراصف التسلسلي (העלמדת רצפים - Sequence Alignment). مُقارنة تسلسلين – أو أكثر – تُمكن من التعرف على المناطق وعلى المواضع التي تتشابه فيها التسلسلات، وكذلك على المناطق والمواضع التي تختلف فيها التسلسلات أيضاً.

לعملية المُقارنة بين التسلسلات أهمية كبيرة في البحث البيوإنفورماتي. المناطق المتشابهة في التسلسلات الموجودة في الأنواع المتشابهة تدلّ عادة على وظيفة و/ أو مبنى مُتشابهين. تُسمى التسلسلات المتشابهة أيضاً باسم تسلسلات هومولوجية. فيما يلي عدّة أمثلة نستخدم فيها الأداة ClustalW لإجراء ترصاف تسلسلات :

- مُقارنة بين تسلسل نوكليوتيدات جين سليم وجين طافر تُؤدّي إلى الكشف عن مكان الطفرة في التسلسل وكذلك عن نوع هذه الطفرة (طفرة استبدال، طفرة نقص وغيرها). بشكل مُشابه بالإمكان المُقارنة بين تسلسل بروتين سليم وتسلسل بروتين طافر والمُقارنة بينهما.
- مُقارنة بين تسلسلات بروتينية مُتشابهة، مثلاً بروتينات تابعة لنفس العائلة البروتينية، تعرض المناطق المتشابهة (المناطق المحفوظة) في هذه البروتينات وأماكنها كما وتعرض المناطق المُختلفة بينها أيضاً. منطقة محفوظة بإمكانها أن تُشير إلى مبنى ووظيفة مُتشابهين.
- مُقارنة بين تسلسل RNA رسول (المكوّن من اكسونات فقط) وبين مقطع جين من الكروموسوم (الجين يضمّ اكسونات وانترونات)، تُشير إلى المكان الدقيق للاكسونات في تسلسل هذا الجين.
- مُقارنة تسلسلات (جينات أو بروتينات) مصدرها من أنواع مُختلفة من الكائنات تُشير إلى مدى التقارب والتشابه بين هذه الأنواع. تشابه كبير بين تسلسلات من نوعين من الكائنات يدلّ على تقارب كبير بين النوعين (קרבה אבולוציונית) (للتوسع اقرأ في المُقدمة " אבולוציה" في فصول المُقدمة).



ClustalW – يُمكن الوصول إلى الأداة بواسطة الرابط:

<http://www.ebi.ac.uk/Tools/clustalw2/index.html>



اهلا بكم في الجولة الإرشادية لأداة ClustalW. ClustalW هو برنامج لمُقارنة التسلسلات، Sequence Alignment.

عندما يكون لدينا تسلسلان أو أكثر فإن الأداة تقوم بوضع التسلسلات الواحد مُقابل الآخر لإيجاد المناطق المُتشابهة والمناطق المُختلفة بين التسلسلات. لمُقارنة التسلسلات أهمية كبيرة من في البحث البيوإنفورماتي. مثلاً، تكشف المُقارنة بين تسلسل من إنسان سليم وتسلسل من إنسان مريض بمرض وراثي عن سبب هذا المرض. في هذا المثال يُمكن أن نرى أن طفرة نقص ثلاثة نوكلوتيدات أدت إلى مرض التليف الكيسي.

مُقارنة تسلسل جين من إنسان سليم (HEALTHY) مع جين من إنسان مريض (SICK)

HEALTHY	GGATTGGGGAAATTAATTTGAGAAAGCAAAACAATAAACAATAGAAAACTTCTAAT	60
SICK	GGATTGGGGAAATTAATTTGAGAAAGCAAAACAATAAACAATAGAAAACTTCTAAT	60

HEALTHY	GGTGATGACAGCCCTTCTTCAGTAATTTCTCACTTCTTGGTACTCCTGTCTGAAAGAT	120
SICK	GGTGATGACAGCCCTTCTTCAGTAATTTCTCACTTCTTGGTACTCCTGTCTGAAAGAT	120

HEALTHY	ATTAAATTTCAAGATAGAAAGAGGACAGTTGTGGCGGTGCTGGATCCACTGGAGCAGGC	180
SICK	ATTAAATTTCAAGATAGAAAGAGGACAGTTGTGGCGGTGCTGGATCCACTGGAGCAGGC	180

HEALTHY	AAGACTTCACTTCTAATGGTGATTATGGGAGAACTGGAGCCTTCAGAGGGTAAAAATTAAG	240
SICK	AAGACTTCACTTCTAATGGTGATTATGGGAGAACTGGAGCCTTCAGAGGGTAAAAATTAAG	240

HEALTHY	CACAGTGAAGAATTTTCATTCTGTCTCAGTTTTCCTGGATTATGCGCTGGCACCATTAAA	300
SICK	CACAGTGAAGAATTTTCATTCTGTCTCAGTTTTCCTGGATTATGCGCTGGCACCATTAAA	300

HEALTHY	GAAAAATATCACTTTGGTGTTTCCCTATGATGAATATAGATACAGAAGCGTCATCAAAGCA	360
SICK	GAAAAATATCACTTTGGTGTTTCCCTATGATGAATATAGATACAGAAGCGTCATCAAAGCA	357

HEALTHY	FTTTCAGAGAAAGACAATATAGTTCTTGGAGAA	420
SICK	FTTTCAGAGAAAGACAATATAGTTCTTGGAGAA	417

HEALTHY	GGTGGAATCACACTGAGTGGAGGTCAACGAGCAAGAAATTTCTTTAGCAAGAGCAGTATAC	480
SICK	GGTGGAATCACACTGAGTGGAGGTCAACGAGCAAGAAATTTCTTTAGCAAGAGCAGTATAC	477

HEALTHY	AAAGATGCTGATTTGTATTTATTAGACTCTCCTTTGGATACCTAGATGTTTTAACAGAA	540
SICK	AAAGATGCTGATTTGTATTTATTAGACTCTCCTTTGGATACCTAGATGTTTTAACAGAA	537

HEALTHY	AAAGAAATATTTGAAAG	557
SICK	AAAGAAATATTTGAAAG	554

מוטציות חסר של 3 נוקלאוטידים היא הגורם

طفرة نقص 3 نوكلوتيدات
سببت مرض التليف الكيسي

واجهة الأداة ClustalW

هذه هي واجهة الأداة ClustalW. في البداية يجب أن نُشير إلى نوع التسلسلات التي نقوم بمقارنتها (تسلسل من الأحماض النووية أو تسلسل من الأحماض الأمينية) عن طريق اختيار protein أو DNA في النافذة المناسبة. نقوم بتزويد نافذة تزويد التسلسلات الموجودة في واجهة الأداة بالتسلسلات التي نرغب بمقارنتها، يجب أن تكون هذه التسلسلات بصيغة FASTA، في هذه الصيغة يبدأ التسلسل بالرمز ">" ويليه اسم التسلسل، بعد ذلك يُعرض تسلسل البروتين أو الجين على شكل تسلسل من الحروف، بحيث أن كل حرف يُشير إلى حامض أميني (في تسلسل البروتين) أو إلى نوكلوتيد (في تسلسل الجين). بعد تزويد التسلسل الأول، نزوّد النافذة بالتسلسل الثاني بنفس الطريقة، وهكذا حتى نزوّد الأداة بجميع التسلسلات التي نرغب بالمقارنة بينها.

The screenshot shows the ClustalW2 web interface with several annotations in Arabic:

- Annotation 1:** "اختيار نوع التسلسلات التي نُقارنها (بروتينات، DNA)" with an arrow pointing to the dropdown menu in the "STEP 1 - Enter your input sequences" section.
- Annotation 2:** "نافذة تزويد التسلسلات التي نُقارن بينها" with an arrow pointing to the text input area in the same section.
- Annotation 3:** "زر تنفيذ المقارنة" with an arrow pointing to the "Submit" button at the bottom of the page.
- Annotation 4:** "حلون لاهوت ركافيم لهشואה" with an arrow pointing to the "Submit" button.

The interface includes a navigation bar with "EMBL-EBI" and "EB-eye Search" logos, a search bar, and a menu with "Databases", "Tools", "EBI Groups", "Training", "Industry", "About Us", and "Help". The main content area is titled "ClustalW2 - Multiple Sequences Alignment" and contains instructions for using the tool, including "STEP 1 - Enter your input sequences", "STEP 2 - Set your Pairwise Alignment Options", and "STEP 3 - Set your Multiple Sequence Alignment Options".

مُقارنة بين تسلسلين

في العرض التوضيحي الأول سنقوم بمُقارنة تسلسلين قصيرين ومُتشابهين جدًا من الأحماض الأمينية. نزود الأداة بالتسلسلات التي نرغب بمُقارنتها، نضغط على الزر Submit وننتظر. عملية المُقارنة يُمكن أن تستمرّ عدّة ثواني.

تضمّ صفحة النتيجة التي نحصل عليها تفاصيل كثيرة، لكننا سنركّز في بطاقة الاختيار Alignment التي تصف المُقارنة بين التسلسلين. في السطرين الأول والثاني، يتمّ ترتيب التسلسلين الواحد فوق الآخر، بحيث يُعرض اسم التسلسل في بداية كل سطر. عندما نضغط على الزر Show Colors تقوم الأداة بتلوين التسلسلات بحسب صفاتها الكيميائية. في السطر الثالث تتمّ الإشارة بالرمز "*" تحت المواضيع التي تتطابق فيها الأحماض الأمينية في التسلسلين. لا تتمّ الإشارة بتأنا تحت المواضيع التي تختلف فيها الأحماض الأمينية بين التسلسلين. من خلال مُقارنة التسلسلات يُمكن أن نرى أنّ الحامض الأميني الخامس فقط يضمّ أحماض أمينية مُختلفة: الحامض الأميني ليوسين (L) في التسلسل الأول والحامض الأميني هستيديين (H) في التسلسل الثاني. يختلف هذان الحامضان عن بعضهما بصفاتها الكيميائية والفيزيائية بشكل كبير جدًا، لذلك لا يحتوي السطر الأخير في هذا الموضوع على أي إشارة.

The screenshot shows the EMBL-EBI ClustalW2 Results page. The main content is a multiple sequence alignment of two sequences, SEQ1 and SEQ2. The alignment is displayed with color coding for amino acids and asterisks (*) indicating matches. The page includes navigation tabs like 'Alignments', 'Result Summary', 'Guide Tree', 'Submission Details', and 'Submit Another Job'. A 'Hide Colors' button is highlighted with a red box. Annotations in Hebrew explain the interface elements:

- زر لعرض / لإخفاء ألوان التسلسلات**: زر لإخفاء / لإظهار ألوان التسلسلات (Hide Colors button)
- رقم الموضع (طول) في كل تسلسل**: رقم الموضع (طول) في كل تسلسل (Position/Length in each sequence)
- اسماء التسلسلات**: اسماء التسلسلات (Sequence names)
- اسم التسلسل**: اسم التسلسل (Sequence name)
- رقم التسلسل**: رقم التسلسل (Sequence number)
- سطر المُقارنة**: سطر المُقارنة (Comparison line)
- النقص بإشارة يدل على أحماض أمينية مُختلفة**: النقص بإشارة يدل على أحماض أمينية مُختلفة (Gap indicates different amino acids)
- الإشارة "*" للإشارة إلى أحماض أمينية مُتطابقة**: الإشارة "*" للإشارة إلى أحماض أمينية مُتطابقة (Asterisk "*" indicates identical amino acids)

نتمتع في مقارنة إضافية. نُزود النافذة بالتسلسلات ونضغط على Submit. في هذه المرة، تُشير نتيجة المقارنة تحت الموضع الخامس بالإشارة ":", هذه الإشارة تدلّ على أنه بالرغم من أنّ الأحماض الأمينية الموجودة في الموضع الخامس في التسلسلين مختلفة، إلا أنّ لها صفات كيميائية وفيزيائية متشابهة جدًا. في هذا المثال الحامض الأميني ليوسين (L) والحامض الأميني ايزوليوسين (I) هي أحماض أمينية هيدروفوبية.

الآن ننظر إلى مقارنة تسلسلات ذات أطوال مختلفة. نتيجة المقارنة تدلّ على أنه في الموضع الثالث والرابع للتسلسل الأول لا يوجد أحماض أمينية تُقابلها في التسلسل الثاني. هذه الإشارة تُسمى باللغة الإنجليزية GAP، ويُشار لها في السطر الثالث بإشارة عارضة "-". معنى هذا أنّ التسلسل الأول أطول من التسلسل الثاني بحامضين أمينيين.

المعنى		الإشارة	
المُقارنة بين تسلسل نوكلونتيديات	المُقارنة بين تسلسل أحماض أمينية	الاسم	الرمز
نوكلونتيديات مُتطابقة في التسلسلين	أحماض أمينية مُتطابقة في التسلسلين	نجمة	"*"
	أحماض أمينية مُختلفة، لكنها ذات <u>صفات كيميائية وفيزيائية مُتشابهة جدًا</u>	نقطتان	":"
	أحماض أمينية مُختلفة، لكنها ذات <u>صفات كيميائية وفيزيائية مُتشابهة قليلاً</u>	نقطة	."
<u>نقص</u> نوكلونتيدي في أحد التسلسلين	<u>نقص</u> حامض أميني في أحد التسلسلين	عارضة	"-"
نوكلونتيديات مُختلفة في التسلسلين	أحماض أمينية مُختلفة في التسلسلين	فراغ	" "

المُقارنة بين عدّة تسلسلات

الأداة ClustalW لا تُلزم المُستخدم بالمُقارنة بين زوج من التسلسلات فقط. عملياً، بالإمكان المُقارنة بين عدد كبير من التسلسلات . من الصعب التمييز بين المناطق المُتشابهة والمناطق المُختلفة عند مُقارنة عدد كبير من التسلسلات. لذلك نستعين بالأداة ClustalW . نستطيع أن نرى في صفحة النتيجة، تراصف التسلسلات والتمييز بسهولة بين المواضع التي تتواجد فيها أحماض أمينية مُتطابقة، مُتشابهة أو مُختلفة في التسلسلات الأربعة المعروضة.

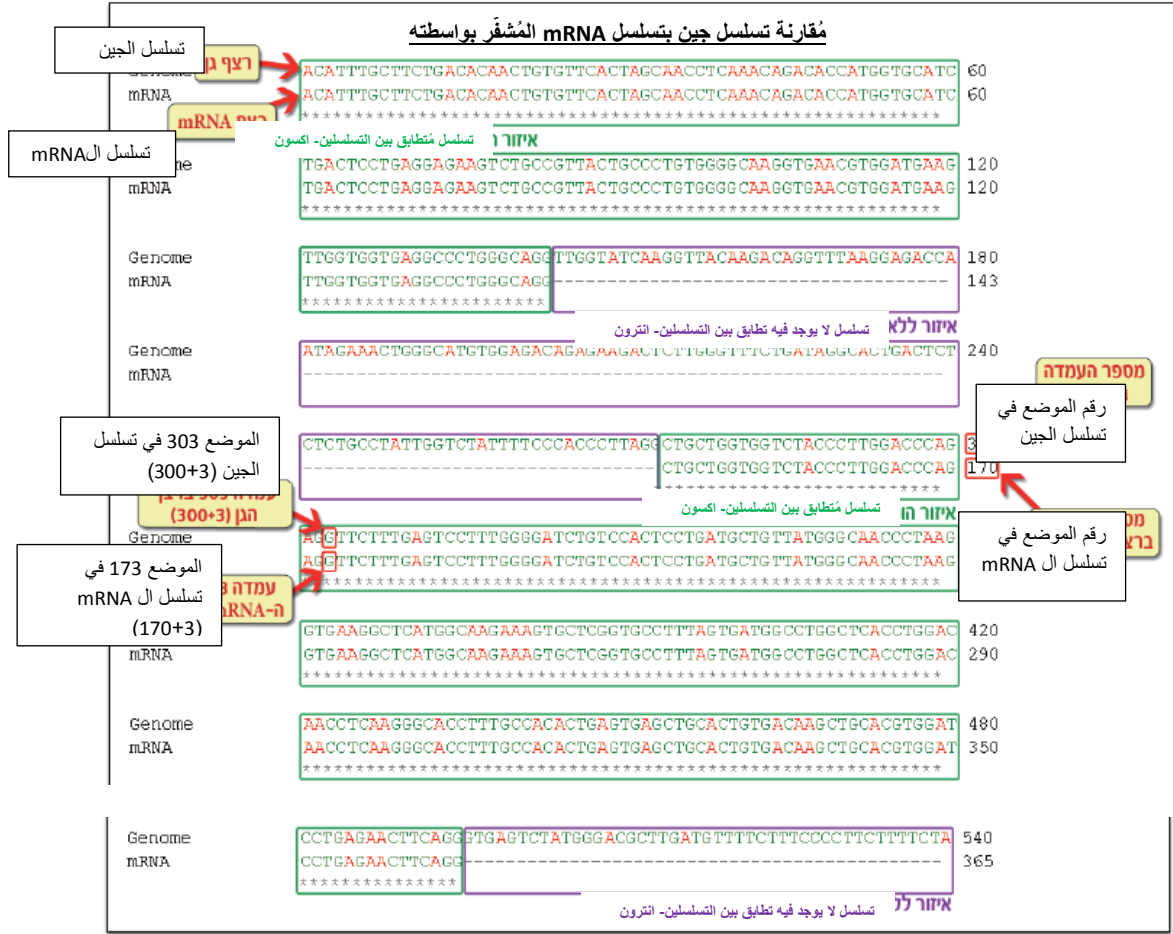
The screenshot shows the ClustalW2 Results page with the following annotations:

- AL**: اسماء التسلسلات (Sequence names)
- ple**: التسلسلات (Sequences)
- رقم الموضع (طول) في كل تسلسل**: Position (length) in each sequence
- شמות הרצפים**: Sequence markers
- سطر المُقارنة**: Comparison line
- شורת המשוואה**: Alignment line
- حماض أميني في التسلسل الأول يُشابه**: Amino acid in the first sequence is similar to the amino acid in the remaining sequences
- حומצה أمينية**: Amino acid
- دומה לאל**: Similar to
- أحماض أمينية مُتطابقة في كل التسلسلات**: Identical amino acids in all sequences
- نقص بحماض أميني في التسلسل الرابع بالمُقارنة مع بقية التسلسلات**: Missing amino acid in the fourth sequence compared to the remaining sequences
- حומצות**: Acidity
- أحماض أمينية مُختلفة في كل التسلسلات**: Different amino acids in all sequences

الأمثلة التي عُرضت حتى الآن ضمّت تسلسلات قصيرة جداً، عادة تكون تسلسلات البروتينات وتسلسلات النوكلونتيديات مكوّنة من تسلسلات أطول بكثير من التسلسلات التي عُرضت في الأمثلة السابقة. مُقارنة مثل هذه التسلسلات وتحليل نتائج المُقارنة بدون استخدام برنامج محوسب هي تقريباً أمر غير مُمكن ويحتاج إلى وقت طويل. بمُساعدة الأداة ClustalW بإمكاننا مُقارنة عدد كبير من التسلسلات بسهولة وسرعة كبيريتين.

أرقام المواضع في التسلسلات

في مُقارنة التسلسلات تظهر أرقام مواضع في نهاية كل سطر من أسطر المُقارنة، ويُعبر هذا الرقم عن الموضع الأخير في السطر، في السطر الأول للمُقارنة يظهر الرقم 60 بجانب كل واحد من التسلسلين. انتبه لا تكون أرقام المواضع مُتطابقة دائماً في التسلسلين. مثلاً، في سطر المُقارنة الخامس، رقم الموضع الأخير في تسلسل الجين هو 300 بينما في تسلسل ال RNA رسول الرقم هو 170. انتبهوا: لحساب رقم موضع مُعين داخل التسلسل، يجب أن نستعين برقم الموضع الأكثر قُرْباً له.



تلخيص

ClustalW هو برنامج لمُقارنة تسلسلات، Sequence Alignment. يُمكن أن نُقارن بواسطة تسلسلات من النوكليوتيدات أو تسلسلات من الأحماض الأمينية. يُساعد برنامج ClustalW على وضع التسلسلات مُقابل بعضها ومُقارنتها، حتى نستطيع أن نُميز بسهولة مواضع الأحماض الأمينية المُتطابقة، المُتشابهة أو المُختلفة في التسلسلات. مناطق مُتشابهة (هومولوجية) في التسلسلات تدل على تشابه في مبنى البروتينات وبالتالي في وظائفها.

أنتم مدعوون لاختبار الأداة والتمرّن على استعمالها، ننمّي أن تستمتعوا بوقتكم. بالنجاح!